

CONTENTS

List of Figures xiii

List of Tables xv

List of Boxes xvii

Acknowledgments xix

1. Introduction 1

The Problem and Objectives 1

Outline of the Book 3

2. Rational Inference: A Constrained Optimization Framework 10

Inference Under Limited Information 11

Qualitative Arguments for Rational Inference 11

Probability Distributions: The Object of Interest 12

Constrained Optimization: A Preliminary Formulation 15

The Basic Questions 18

Motivating Axioms for Inference Under Limited Information 19

Axioms Set A: Defined on the Decision Function 20

Axioms Set B: Defined on the Inference Itself 20

Axioms Set C: Defined on the Inference Itself 21

Axioms Set D: Symmetry 22

Inference for Repeated Experiments 22

Axioms Versus Properties 24

3. The Metrics of Info-Metrics 32

Information, Probabilities, and Entropy 32

Information Fundamentals 32

Information and Probabilities 37

Information and Entropy 39

Information Gain and Multiple Information Sources 43

Basic Relationships 43

Entropy and the Grouping Property 44

Relative Entropy 46

Mutual Information 47

Axioms and Properties 49

Shannon's Axioms 49

Properties 49

| | |
|--|-----|
| 4. Entropy Maximization | 59 |
| Formulation and Solution: The Basic Framework | 60 |
| <i>Information, Model, and Solution: The Linear Constraints Case</i> | 60 |
| Model Specification | 60 |
| The Method of Lagrange Multipliers: A Simple Derivation | 61 |
| Information, Model, and Solution: The Generalized Constraints Case | 68 |
| <i>Basic Properties of the Maximal Entropy Distribution</i> | 71 |
| Discussion | 72 |
| Uniformity, Uncertainty, and the Solution | 72 |
| Conjugate Variables | 74 |
| Lagrange Multipliers and Information | 76 |
| The Concentrated Framework | 79 |
| Examples in an Ideal Setting | 82 |
| Geometric Moment Information | 82 |
| Arithmetic Moment Information | 83 |
| Joint Scale and Scale-Free Moment Information | 87 |
| Likelihood, Information, and Maximum Entropy: A Qualitative Discussion | 87 |
| 5. Inference in the Real World | 107 |
| Single-Parameter Problems | 108 |
| Exponential Distributions and Scales | 108 |
| Distribution of Rainfall | 108 |
| The Barometric Formula | 110 |
| Power and Pareto Laws: Scale-Free Distributions | 112 |
| Distribution of Gross Domestic Products | 113 |
| Multi-Parameter Problems | 114 |
| Size Distribution: An Industry Simulation | 114 |
| Incorporating Inequalities: Portfolio Allocation | 117 |
| Ecological Networks | 122 |
| Background | 123 |
| A Simple Info-Metrics Model | 124 |
| Efficient Network Aggregation | 126 |
| 6. Advanced Inference in the Real World | 135 |
| Interval Information | 136 |
| Theory | 136 |
| Conjugate Variables | 139 |
| Weather Pattern Analysis: The Case of New York City | 140 |
| Treatment Decision for Learning Disabilities | 143 |
| Background Information and Inferential Model | 143 |
| A Simulated Example | 145 |
| Brain Cancer: Analysis and Diagnostics | 147 |
| The Information | 148 |

| | |
|---|-----|
| <i>The Surprisal</i> | 151 |
| <i>Bayesian Updating: Individual Probabilities</i> | 154 |
| 7. Efficiency, Sufficiency, and Optimality | 165 |
| Basic Properties | 166 |
| Optimality | 166 |
| Implications of Small Variations | 167 |
| Efficiency | 169 |
| Statistical Efficiency | 169 |
| Computational Efficiency | 175 |
| Sufficiency | 176 |
| Concentration Theorem | 178 |
| Conditional Limit Theorem | 180 |
| Information Compression | 180 |
| 8. Prior Information | 194 |
| A Preliminary Definition | 195 |
| Entropy Deficiency: Minimum Cross Entropy | 196 |
| Grouping Property | 200 |
| Surprisal Analysis | 209 |
| Formulation | 209 |
| Extension: Unknown Expected Values or Dependent Variables | 211 |
| Transformation Groups | 211 |
| The Basics | 212 |
| Simple Examples | 215 |
| Maximum Entropy Priors | 221 |
| Empirical Priors | 221 |
| Priors, Treatment Effect, and Propensity Score Functions | 222 |
| 9. A Complete Info-Metrics Framework | 231 |
| Information, Uncertainty, and Noise | 232 |
| Formulation and Solution | 234 |
| A Simple Example with Noisy Constraints | 242 |
| The Concentrated Framework | 245 |
| A Framework for Inferring Theories and Consistent Models | 249 |
| Examples in an Uncertain Setting | 250 |
| Theory Uncertainty and Approximate Theory: Markov Process | 250 |
| Example: Mixed Models in a Non-Ideal Setting | 254 |
| Uncertainty | 259 |
| The Optimal Solution | 259 |
| Lagrange Multipliers | 261 |
| The Stochastic Constraints | 262 |
| The Support Space | 262 |
| The Cost of Accommodating Uncertainty | 264 |

| | |
|---|------------|
| Visual Representation of the Info-Metrics Framework | 264 |
| Adding Priors | 268 |
| 10. Modeling and Theories | 281 |
| Core Questions | 282 |
| Basic Building Blocks | 284 |
| <i>Problem and Entities</i> | 284 |
| <i>Information and Constraints</i> | 285 |
| <i>Incorporating Priors</i> | 286 |
| <i>Validation and Falsification</i> | 286 |
| <i>Prediction</i> | 288 |
| A Detailed Social Science Example | 288 |
| <i>Characterizing the Problem</i> | 288 |
| <i>Introducing the Basic Entities</i> | 289 |
| <i>Information and Constraints</i> | 291 |
| Production | 291 |
| Consumption | 292 |
| Supply and Demand | 292 |
| Individual Preferences | 292 |
| Budget Constraints | 293 |
| <i>The Statistical Equilibrium</i> | 294 |
| <i>Economic Entropy: Concentrated Model</i> | 296 |
| <i>Prices, Lagrange Multipliers, and Preferences</i> | 297 |
| <i>Priors, Validation, and Prediction</i> | 298 |
| <i>Model Summary</i> | 299 |
| Other Classical Examples | 300 |
| 11. Causal Inference via Constraint Satisfaction | 307 |
| Definitions | 308 |
| Info-Metrics and Nonmonotonic Reasoning | 309 |
| <i>Nonmonotonic Reasoning and Grouping</i> | 314 |
| <i>Typicality and Info-Metrics</i> | 316 |
| <i>The Principle of Causation</i> | 316 |
| Info-Metrics and Causal Inference | 318 |
| Causality, Inference, and Markov Transition Probabilities: An Example | 319 |
| <i>The Model</i> | 320 |
| <i>Inferred Causal Influence</i> | 322 |
| 12. Info-Metrics and Statistical Inference: Discrete Problems | 334 |
| Discrete Choice Models: Statement of the Problem | 335 |
| <i>Example: A Die and Discrete Choice Models</i> | 335 |
| Definitions and Problem Specification | 339 |
| The Unconstrained Model as a Maximum Likelihood | 340 |

| | |
|---|------------|
| The Constrained Optimization Model | 341 |
| The Info-Metrics Framework: A Generalized Likelihood | 343 |
| Real-World Examples | 345 |
| <i>Tailoring Political Messages and Testing the Impact of Negative Messages</i> | 345 |
| Background on the Congressional Race and the Survey | 346 |
| Inference, Prediction, and the Effect of Different Messages | 346 |
| <i>Is There Racial Discrimination in Home Mortgage Lending?</i> | 347 |
| Background on Loans, Minorities, and Sample Size | 347 |
| Inference, Marginal Effects, Prediction, and Discrimination | 348 |
| The Benefits of Info-Metrics for Inference in Discrete Choice Problems | 351 |
| 13. Info-Metrics and Statistical Inference: Continuous Problems | 357 |
| Continuous Regression Models: Statement of the Problem | 358 |
| Definitions and Problem Specification | 359 |
| Unconstrained Models in Traditional Inference | 359 |
| Rethinking the Problem as a Constrained Optimization | 361 |
| <i>A Basic Model</i> | 361 |
| <i>A General Information-Theoretic Model</i> | 363 |
| Generalized Entropies | 364 |
| Information-Theoretic Methods of Inference: Zero-Moment Conditions | 367 |
| <i>Specific Cases: Empirical and Euclidean Likelihoods</i> | 367 |
| Exploring a Power Law: Shannon Entropy Versus Empirical Likelihood | 371 |
| <i>Theoretical and Empirical Examples</i> | 373 |
| Information-Theoretic Methods of Inference: Stochastic Moment Conditions | 376 |
| <i>The Support Spaces</i> | 380 |
| <i>A Simulated Example</i> | 381 |
| Misspecification | 386 |
| The Benefits of Info-Metrics for Inference in Continuous Problems | 388 |
| Information and Model Comparison | 390 |
| 14. New Applications Across Disciplines | 411 |
| Option Pricing | 412 |
| <i>Simple Case: One Option</i> | 413 |
| <i>Generalized Case: Inferring the Equilibrium Distribution</i> | 416 |
| <i>Implications and Significance</i> | 418 |
| Predicting Coronary Artery Disease | 418 |
| <i>Data and Definitions</i> | 419 |
| <i>Analyses and Results</i> | 420 |
| <i>The Complete Sample</i> | 420 |

| | |
|--|-----|
| Out-of-Sample Prediction | 423 |
| Sensitivity Analysis and Simulated Scenarios | 424 |
| <i>Implications and Significance</i> | 425 |
| Improved Election Prediction Using Priors on Individuals | 426 |
| <i>Analyses and Results</i> | 427 |
| The Data | 427 |
| The Priors and Analyses | 428 |
| <i>Implications and Significance</i> | 431 |
| Predicting Dose Effect: Drug-Induced Liver Injury | 432 |
| Medical Background and Objective | 433 |
| Data and Definitions | 434 |
| Inference and Predictions | 434 |
| A Linear Model | 434 |
| Analyzing the Residuals: Extreme Events | 437 |
| <i>Implications and Significance</i> | 439 |
| Epilogue | 446 |
| List of Symbols | 449 |
| Index | 451 |